Advance Science Letter 2015

by Dewi Tamara

RESEARCH ARTICLE

Advanced Science Letters
Vol. 21, No 4 page: 826-829, April 2015

Indonesian Mutual Funds Classification Using Clustering Method

Ir. Dewi Tamara MM., MS [1], Shintia Revina, SPd., MSc [2]

[1] Binus Business School - Bina Nusantara University (Email: dtamara@binus.edu)
[2] School of Computer Science - Bina Nusantara University (Email: srevina@binus.edu)

Abstract: *Mutual funds have existed since 1990 as an alternative investment in Indonesia. The objective of this research is to examine the existing classification of mutual funds database. The data of mutual funds is taken from Bloomberg through Portal Reksadana 2013 which covered 690 mutual funds. The existing classification consists of mutual funds fixed income (reksadana pendapatan tetap), equity (reksadana saham), money market (reksadana pasar uang) and structured (reksadana campuran). The existing financial attributes consists of the net asset value, percentage annualized return the last 6 months, 1 year, 3 years, 5 years and year-to-date. This paper uses K-means clustering to propose new classification of Indonesian mutual funds. The result reveals that mutual funds in equity and fixed income belong to its group. However, mutual funds money market is belong to mutual fund fixed income and mutual funds structures are identified to mutual funds equity. Furthermore, we find that in average 43% of Indonesian mutual funds are misclassified in accordance with their attributes. Finally, it is suggested to re-group the mutual funds into smaller classification, which has lower rates of misclassified mutual funds and possibility to achieve better performances in terms of its percentage annualized return.*

Key words: *mutual fund, reksadana, clustering analysis.*

1. INTRODUCTION

A mutual fund is an alternative investment which pools money from shareholders and invests in a diversified portfolio of securities. Mutual funds are usually classified based on their investment objectives or asset classes.

In Indonesia, mutual funds began as an alternative investment since 1990 when closed-end funds were established based on Keputusan Menteri Keuangan 1548 (Act no. 1548 Minister of Finance). Furthermore, in 1995, Law no.8 on the capital market allowed both open-end funds and closed-ends fund to operate in Indonesia.

According to Danila (2012), the amount of Net Asset Value, indication of expansion of mutual funds improved from Rp 8 trillion in 2001 to Rp 114.37 trillion in 2009. In November 2013, the Net Asset Value reached Rp 197,87 trillion.

The importance of the research about mutual funds performance and classification is growing over the years and countries. However, Indonesian academic research has generally lagged behind investor's interest in mutual fund literature. Hence, since the Net Asset Value improved significantly, the performance and classification of mutual funds in Indonesia is worth to explore.

There are several advantages of having new classification of mutual funds. First, it can be used to develop benchmark for evaluating the performance of mutual funds. Secondly, other database can be built using this classification scheme. Thirdly, it can reduce the misclassification of existing database and finally it can reduce the misclassification and inefficiency of portfolio diversification.

Our study is different with others do. For example, Das (2003) use K-means Clustering method to classify mutual fund in United States. Ours is close to Marathe and Shawky (1999) which categorized mutual funds in Spain using clusters. This paper attempts to investigate whether classification in mutual funds in Indonesia is already achieved its performance objectives. The authors use cluster analysis approach to classify the mutual funds.

The structure paper is consisted of Section 2, which describes the Literature Review, Section 3 is the Data and Methodology, Section 4 is the Discussion and Section 5 is the Conclusion.

2. MUTUAL FUNDS CLASSIFICATION

The strand of mutual funds scholar can be divided into

two, performance and classification. The performance literature of mutual funds investigates the flow of mutual fund investors as a response to past performance. Ippolito (1992), Chevalier & Ellison (1997), Sirri & Tufano (1998), among others, found that the relation of flow and performance was to be convex. It is explained that high performance mutual funds will receive overproportional inflows, while low performance mutual funds will receive only mild outflows. Gruber (1996) also described how smart money flowing into mutual funds moves in the direction of best performers. In German mutual fund market, Ber, Kempf and Ruenzi (2007) and Jank & Wedow (2010) confirm this convex flow-performance relationship.

The second strand of mutual funds is the classification method. The problem of consistency of funds performance and its stated objective also attracted scholar. Sharpe developed his seminal work in 1992, determining the mutual fund styles applying an asset class factor model (twelve asset classes are employed). In such model each factor represents the return on an asset class and the sensitivities (estimated parameters) are required to sum 1. This is employed to provide a natural method for constructing benchmarks, through a portfolio compounded by a mix of asset classes with the same estimated style.

In 1997, diBartolomeo and Witkowski (1997) classified the mutual funds based on Sharpe (1992) ratio. Their result shows that 40% of equity funds are misclassified. Further investigation using Monte Carlo simulations, they find that misclassification has a significant impact on the investor's ability to diversify his portfolio of mutual funds. They conclude that latter outperforms the official classification, in terms of Sharpe's ratio. From the Spanis case, Matallin and Fernandez (1996) find that the obtained styles from the Sharpe's asset class factor (1992) is in agreement with the official classification, which was established by CNMV.

In this study, we aim to evaluate the mutual fund classification in Indonesia based on their objectives. We use cluster analysis to classify the mutual funds.

2.1 Mutual Funds Industry in Indonesia

According to Danila (2012), Bapepam-LK, as an authorized institution to monitor and supervise the capital market, classifies mutual funds in Indonesia as follows: conventional funds, structured funds, exchange traded funds, and Shari'ah funds. Conventional funds consist of fixed income funds, growth/equity funds, money market funds and balanced funds. Chan (1999) examines a style of fund is more likely to choose stock that outperformed the market index as a benchmark.

Structured Funds is the new funds created by Bapepam-LK that considers low risk in volatile financial market and new variety product funds. These types of funds consist of capital protected funds, guaranteed funds and index funds. The capital protected funds are fixed-income funds with additional features such as initial capital protection and locked-up period of redemption. The allocation is 80 to 90% of NAV will be invested in fixed income instruments such as zero coupon bonds and government bonds and the 20-10% will be invested in equity or derivatives. Guaranteed Fund is a fund which insured investors' capital, usually financial institution. Index Funds are funds that replicate the performance of broad market index, such as Index Harga Saham Gabungan (IHSG), and Jakarta Islamic Index (JII). The index fund is managed passively.

Exchange Traded Funds is the type of funds that similar to index funds, both of them hold investment portfolios that replicate a performance of market index and both of them are passively managed investment instruments. The difference between index funds is the ETF is traded through the stock exchange (Investment company Fact book, 2007), and index fund traded through broker or dealer or fund manager.

Shari'ah Fund is the fund based on Islamic rules, that is the money that invested in mutual funds should be according to Islamic rules. For example, beer industry or conventional banking that applied interest rate.

2.2 Mutual Fund Classification Methods

Mutual fund classification literatures show that investment objective has been widely used as the base of the classification schemes. McDonald (1974) examined the overall performance of a sample of mutual funds relative to their self-declared objective and found a positive relationship between investment objectives and measures of risk. Martin, Keown and Farrell (1982) examined a sample of mutual funds representing five investment objectives and found definite differences in the variability of the funds in alternative classification.

In the other study, Marathe and Shawsky (1999) state that mutual fund managers have several incentives to misclassify their fund's investment objective. First, the managers will get compensated based on their fund's performance and ranking. Higher ranking meant higher performance and the managers has an incentive to intentionally misclassify the fund's objective. They find that a 43% of the mutual funds do not belong to their stated categories and that in many instances self-declared categories of mutual funds are indistinguishable from one another when their classification is based on financial characteristics.

Using the same principle, Kim et al (2000) classify mutual fund using observable attributes (characteristics and investment style) in addition to the risk and return measures. They implement a discriminant analysis. They find that the stated objectives of over 50% of the funds do not match their attributes-based objectives, and the stated objectives of over 33% of the funds depart severely from their stated objectives. Moreover, to classify mutual funds which pick up all nonlinear patterns and relations between the whole of funds attributes. Moreno, Olmeda, and Marco (2000) apply what so called a nonlinear methodology. Using the methodology of Self-Organizing Maps, K-nearest neighbors and the K-means algorithm (clustering method), they reveal that 40% of mutual funds are misclassified in which the funds do not conform to the investment objectives stated in their official category.

2.3 Classification Using Clustering Method

Clustering analysis has been widely used as a classification tool. It is a statistical method of classification which reveals structure and relations in data. The results of this analysis can give direct contribution to the classification process.

In the clustering analysis, the data provided should be displayed in terms of their properties such as attributes and class membership. However, the common problem with the collected data, especially in finance, is the lack of homogeneity among the attributes where the range value of one particular attribute is much greater than that of the other attributes. The attributes with a larger range value will have opportunity to contribute more in the classification process. Therefore, the

attributes should be standardized or normalized to avoid the dominancy of one particular attribute.

In this study, we use K-means clustering method as one the widely used clustering methods. This method is used to classify mutual funds based on the attributes discussed in the next section. In K-means method, the closest cluster center for each mutual fund will be identified and the mutual fund to that cluster will be assigned. This method then re-computes each cluster center based on the mutual funds assigned to that cluster. The assignment of mutual funds to clusters is done iteratively until no more assignments.

As decribed by MacQueen (1967), the K-Means method can be described in the following steps:

Step 1 - Divide the data items into initial clusters

Step 2 - Assign the data items to the cluster with nearest centroid. The centroid in each cluster is the point to which the sum of distances from all data items within its cluster is minimized.

Step 3 - Recalculate the centroid for the cluster receiving or losing the items

Step 4 - Do step 2 and 3 until no more reassignment

K-means clustering method observes the data item as an object in space with specific location. The data items are partitioned in *k* clusters so that the objects in a cluster will be close each other, and in distance with member from another cluster.

3. DATA COLLECTION & METHODOLOGY

3.1 Data Collection

Our database is obtained from Bloomberg through Portal Reksadana 2013. It consists of 690 mutual funds and classifies into for different categories. First category is Reksadana Campuran (Structured Funds), second is Reksadana Saham (Equity Funds), third is Reksadana Pasar Uang (Money Market Funds) and and the last is Reksadana Fixed (Fixed Funds). In Dan (2012), the funds are categorized as conventional funds.

Based on the availability and completeness of the data we selected 6 financial variables to be performed in the cluster analysis. These financial variables consist of Price, Year-To-Date, % return in 6 months, % return in 1 year, % return in 3 years, % return in 5 years. There are 252 funds from 690 that have complete data of above financial variables. These include 75 Reksadana Campuran, 66 Reksadana Saham, 28 Reksadana Pasar Uang and 83 Reksadana Fixed.

3.2 Methodology

All variables in this study are normalized in the pre-processing stage since the data used for this clustering analysis have different unit of measurement (i.e. Price in IDR and YTD in %). The normalization procedure applied equalized the range of all of data to [0, 1]. Normalizing the data gives opportunity to all variables to equally contribute in the clustering process.

3.2.1 K-Means Clustering Analysis

We evaluate the classification of Indonesian mutual funds based on existing database. We use K-Means cluster analysis to classify the funds and compare the result with the investment strategy-based classification. The K-means clustering method classified the funds based on their financial attributes. Thus, the result will show which funds categorized as correctly or incorrectly clustered instances. This could explain why some funds do not perform as their stated objectives.

*EMAIL ADDRESS OF AUTHOR TO WHOM

In determining the cluster center (in terms of distance measure), we use Euclidean distance. Euclidean distance between two objects x and y can be calculated as follows:

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (a_i(x) - a_i(y))^2}$$

where $a_i(x)$ denotes the value of the i-th variable A_i of x.

The closest centroid for each fund then will be identified and the K-means method will assign the fund to that cluster. This is an iterative process until no more assignments can be done.

3.2.2 Tool

We execute the analysis using Weka, software that provides complete data mining tools and algorithms, including K-means clustering method.

4. RESULT

Before we fed the classifier with the data, we normalized all the variables, so they can give equal contribution in the classification process. This is required in our data preprocessing as the variables have different unit of measurement. Table 1 shows the range of each variable before we normalized:

Table 1. Range of Each Financial Variable Before Normalization

Financial Variables	Range
Price	[648, 61808]
Return 6 Month	[-21.92, 23.03]
Return 1 Year	[-24.74, 61.06]
Return 3 Year	[-16.54, 25.14]
Return 5 Year	[-4.81, 42.04]
Return YTD	[-4.94, 7.17

The result of the normalization applied gives an equalized the range of all of attributes to [0, 1]. Now the mutual fund data is ready to be clustered. Thus, the attributes result in different range as shown above.

4.1 Cluster Centroids

The result shows that the 5 iterations are needed until no more reassignments take place. The cluster centroids are shown in table 2 below:

Table 2. Clusters Centroid

Financial Variables	All Data	Cluster 1*	Cluster 2*
Price	0.0436	0.0764	0.0206
Return 6 Months	0.4449	0.4067	0.4718
Return 1 Year	0.2659	0.2917	0.2478
Return 3 Years	0.5043	0.5588	0.4660
Return 5 Years	0.3964	0.5983	0.2546
Return YTD	0.5182	0.6547	0.4222

**Cluster 1: Saham*
Cluster 2: Fixed

From the above result, instead of classifying the funds into four categories as stated in the investment strategy, the funds now are classified into two clusters, namely cluster Saham and Fixed.

4.2 Funds Classification

Table 3 shows the fund classification based on the financial variables using clustering method:

Table 3. Mutual Funds Classification Assigned to New Cluster

Types of Fund	Assigned to Cluster	
	Cluster 1	Cluster 2
Reksadana Campuran	41	34
Reksadana Saham	62	4
Reksadana Pasar Uang	0	28
Reksadana fixed	1	82

*Incorrectly clustered instances: 42.8571%

From the above result, based on its financial characteristics, all Reksadana Pasar Uang are identified as Reksadana Fixed. Whereas, most of Mutual Funds Structure are identified as Mutual Funds Equity (Reksadana saham). It is also revealed that the number of incorrectly clustered funds is 42.8571%.

5. Discussion & Conclusion

Indonesian mutual funds are classified using the investment objectives. There four classification of the mutual funds based on investment objectives which are mutual funds fixed income (reksadana pendapatan tetap), mutual funds equity (reksadana saham), mutual funds money market (reksadana pasar uang), and mutual funds structured (reksadana campuran). The objective of this paper is to investigate the existing mutual fund classification in Indonesia.

In this paper, mutual funds in the database are classified using cluster analysis. The cluster analysis analyzed the attributes of mutual funds that consist of percentage of annualized return from 6 months to year-to-date. The attributes in the mutual funds are normalized to uniform scale of measurement.

Using 5 times iterations, the current four mutual funds classification is reduced to two new classifications which are mutual funds fixed income and equity. It is found that in average 43% of mutual funds are misclassified.

The use of statistical techniques to re-classify the existing mutual funds categories instead of investment objectives has the potential of better explaining the differences in the return performance of the various funds. The analysis found an indication that the various mutual funds seem to behave similarly using k-means clustering. Using cluster, the current mutual funds could be efficiently divided into two groups. These two groups suggested are the fixed income where mutual funds money market show similar behavior, and the mutual funds equity where the mutual funds structure clustered in the same group. The return-risk characteristics are too alike to justify determining different categories. Further analysis needs to be done to investigate other attributes might be presented as mutual funds performance such as incentive fee and liquidity of the mutual funds.

REFERENCES

Babinec Tony (2002): "Editoral: Some musings on Cluster Analysis, "Journal of Targeting, Measurement and Analysis for Marketing, Vol. 10, No. 4, p 301-303.

Danila, N., (2012), Estimating the Risk of Mutual Funds in Indonesia by Employing Value at Risk (VaR), Asian Journal of Business and Accounting 5(2).

Das, N., (2003), Mutual Fund Classification using K-means Clustering Method, 9th International Conference on computing in Economics and Finance, University of Washington, Seattle.

Johnson Richard A. and Wichern Dean W. (1998): "Applied Multivariate Statistical Analysis," *Prentice Hall*

Kuo R. J., Ho L. M. and Hu C. M. (2002): "Cluster Analysis in Industrial Market Segmentation through Artifical Neural Network," Computers and Industrial Engineering, Vol. 42, p 391-399.

Marathe, A., and H.A. Shawky, (1999), Categorizing Mutual Funds Using Clusters. Advances in Quantitative analysis of Finance and Accounting, 1, vol 7, 199-204.

Moreno, D., I. Olmeda, and P. Marco., (2000), A New Look on Classification of Spanish Mutual Funds, Working Paper Draft.

Received: .. May 2014. Accepted: … July 2014

Advance Science Letter 2015

ORIGINALITY REPORT

33%	30%	9%	10%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.ivie.es Internet Source	8%
2	ejournal.um.edu.my Internet Source	7%
3	pdfs.semanticscholar.org Internet Source	5%
4	www.c3.lanl.gov Internet Source	4%
5	www.econstor.eu Internet Source	1%
6	Submitted to iGroup Student Paper	1%
7	David Moreno, Paulina Marco, Ignacio Olmeda. "Self-organizing maps could improve the classification of Spanish mutual funds", European Journal of Operational Research, 2006 Publication	1%
	www.sunlife.com.ph	

8	Internet Source	1%
9	www.nationalbank.at Internet Source	1%
10	Submitted to National University of Singapore Student Paper	1%
11	Sakakibara, Takumasa, Tohgoroh Matsui, Atsuko Mutoh, and Nobuhiro Inuzuka. "Clustering Mutual Funds Based on Investment Similarity", Procedia Computer Science, 2015. Publication	1%
12	Qiang Bu, Nelson Lacey. "What drives mutual fund flows: profit or panic?", Review of Accounting and Finance, 2013 Publication	1%
13	mjs.um.edu.my Internet Source	<1%
14	xiforofinanzas.ua.es Internet Source	<1%
15	www.wifo.ac.at Internet Source	<1%
16	cyberleninka.ru Internet Source	<1%
17	Submitted to University of Birmingham Student Paper	<1%

18	e-archivo.uc3m.es Internet Source	<1%
19	www.wellian-is.com Internet Source	<1%
20	open.uct.ac.za Internet Source	<1%
21	academic.oup.com Internet Source	<1%

Exclude quotes	On	Exclude matches	< 10 words
Exclude bibliography	On		